Current status and issues of traffic light recognition technology in Autonomous Driving System

Naoki Suganuma*, non-member and Keisuke Yoneda†, non-member

SUMMARY. Autonomous driving technology is currently attracting a lot of attention as a technology that will play a role in the next generation of mobility. For autonomous driving in urban areas, it is necessary to recognize various information. Especially, the recognition of traffic lights is important in crossing intersections. In this paper, traffic light recognition technology developed by the authors was evaluated using onboard sensor data during autonomous driving in the Tokyo waterfront area as an example of traffic light recognition technology. Based on the results, it was found that traffic lights could be recognized with an accuracy of approximately 99% to carry out the decision making for intersection approaching. However, from the evaluation results, it was also confirmed that traffic light recognition became difficult under situations involving occlusion by other object, background assimilation, nighttime conditions, and backlight by sunlight. It was also confirmed that these effects are mostly temporary, and do not significantly affect decision-making to enter intersections as a result of utilizing information from multiple traffic lights installed at an intersection. On the other hand, it is expected that recognition with current onboard cameras will become technically difficult during situations in which not all traffic lights are visually recognizable due to the effects of back or front light by sunlight when stopped at the stop line of an intersection. This paper summarizes these results and presents the necessity of appropriate traffic light installation on the assumption of recognition by onboard cameras.

key words: Traffic light recognition, Autonomous driving, image processing, SIP-adus.

1. Introduction

Autonomous driving technology is currently attracting a lot of attention as a technology that will play a role in the next generation of mobility. Research on autonomous vehicles has been conducted since around 1960 and has a long history in various countries around the world. At the beginning of the development of research, most of the research was aimed at autonomous driving on dedicated roads such as expressways, but autonomous driving systems, which have attracted attention in recent years, have become possible to autonomous driving on various roads, including urban area [1-3].

For autonomous driving in urban areas, it is necessary to recognize various information. Especially, the recognition of the traffic light is important when crossing the intersections. In traffic light recognition, it is necessary to reliably acquire the current information of the traffic light such as emitted color of the signal light and the direction of the arrow signal. This information is important from the viewpoint of safety because misidentifying traffic lights not only leads to traffic rule violations, but directly causes traffic accidents.

A common means of acquiring current information from traffic lights with onboard devices is analyzing images obtained from cameras [1,2]. The acquisition of current information through communication with infrastructure (V2I) has also been considered. Although the former enables autonomous driving even in regular intersections without wireless infrastructure, there are situations in which reading the status of traffic lights from images becomes difficult under special environmental conditions such as backlight by sunlight. However, if communication can be established, the latter has advantages such as being able to acquire current traffic light information without being affected by weather and other factors, making it desirable to develop it as soon as possible. However, there are disadvantages in that autonomous driving is difficult in environments lacking communication infrastructure, and infrastructure development is costly. Therefore, it may be necessary to ascertain the limitations of traffic light recognition by image recognition and situations in which recognition becomes difficult and explore the environments in which the installation of traffic lights with wireless infrastructure should be prioritized.

In Japan, SIP-adus (Cross-ministerial Strategic Innovation Promotion Program - Automated Driving for Universal Services) project, a research and development project related to autonomous driving, is currently carried out as a national project. As part of this project, our research group is studying issues related to traffic light recognition. In this paper, the problems and future prospects of traffic light recognition are presented referring to the traffic light recognition technology developed by the authors as an example.

†The authors are with Kanazawa University, Kakuma-machi, Kanazawa, Ishikawa, 920-1192 Japan.

a) E-mail: suganuma@staff.kanazawa-u.ac.jp
2. Autonomous Driving Vehicle and Tokyo Waterfront Area Demonstration

Fig.1 shows an autonomous driving vehicle developed by the authors. Our research group has been developing various technologies related to autonomous driving since around 1998, from methods of generating and updating the HD maps required for automatic driving, to recognition technology based on onboard sensors and path planning technology that realizes autonomous driving in urban areas. The test vehicle shown in Fig.1 is equipped with a fully original autonomous driving system including middleware developed in our laboratory. Presently, we are using this vehicle to perform test runs of autonomous driving technology on regular public roads in various regions in Japan.

![Fig.1 Autonomous driving vehicle developed by authors.](image)

In the SIP-adus project mentioned above, various demonstration tests for autonomous driving systems have been carried out in the Tokyo waterfront area of Japan, and in our group, test runs has started since September 2019. In the period until March 2021, our autonomous vehicle has been running at a distance of about 2,138km in an autonomous driving mode, and the driving test is still ongoing.

This vehicle is equipped with five LiDARs, eight cameras, and nine-millimeter wave radars as surrounding environmental recognition sensors. For the traffic light-recognizing cameras, in addition to the previously mentioned eight cameras, a two onboard cameras with Full-HD resolution and equipped with an LFM-HDR (LED Flicker Mitigation-High Dynamic Range) function that can acquire images with a wide dynamic range while suppressing flicker generated when shooting LED traffic lights are also installed. Using these cameras, it is possible to recognize traffic lights in various lighting environments. Additionally, using multiple cameras with different field of view, long-distance traffic light recognition is realized, as well as wide angle traffic light recognition. In addition to the cameras for traffic light recognition, an antenna for infrastructure communication is also mounted on the upper part of the vehicle. The Tokyo waterfront area where the test runs took place has the infrastructure to supplement camera-based image recognition with wireless communication-based acquisition of traffic light information, including current traffic light information and traffic light seconds remaining information.

Furthermore, a tightly coupled GNSS/INS navigation system is also installed to estimate the vehicle position and the motion status. In addition to this GNSS/INS, map matching with the above-mentioned surrounding environment recognition sensors and HD maps makes it possible to estimate vehicle positions with an accuracy of about 0.1 to 0.2m in various environments [1,4,5]. Using these sensors and HD maps enables autonomous driving in urban areas.

The next chapter outlines the traffic light recognition technology used in these autonomous driving experiment.

3. Overview of Traffic Light Detection Algorithm

3.1 Requirements for Traffic Light Detection

In this section, we focus on traffic light recognition algorithm using HD maps. There are many studies on traffic light recognition by using vision-sensors. Although HD maps have the information of traffic light positions, the vehicle must recognize the current state of traffic lights in real time because it changes dynamically. For safety deceleration, it is necessary to recognize the current state of traffic lights at distances over 100 m. The required recognition distance can be estimated by calculating the braking distance from the vehicle to the stop line after smoothly recognizing the traffic light state. In studies on vehicle control [6,7], the deceleration without discomfort to passengers is approximately 0.1 G ( =0.098 m/s²). For example, when the vehicle decelerates by 0.1 G while traveling at a velocity of 50 km/h, the braking distance is approximately 98 m. Furthermore, the recognition distance may increase further when considering the case where the traffic light is located at a position away from the stop line. Recognizing traffic lights in the ranges exceeding 100 m is required to make a natural intersection approach in automated driving. In order to implement a practical method of traffic light recognition, it is necessary to discuss the effectiveness of the methods, considering the trade-off between the required performance and the hardware specification. For example, installing a high-resolution camera or a telephoto lens is an easy solution to increase the recognition distance. However, increasing the resolution may increase the processing time. In addition, the field of view is narrowed and traffic lights may be left out. From the point of view of implementing a recognition method, it is important how to recognize small pixel objects. On the other hand, in automated driving using HD maps, the self-localization module precisely estimates the vehicle pose by
map-matching using a range sensor or image sensor [4, 5]. Generally, position accuracy of approximately 0.1 to 0.2 m is considered to be necessary for decision-making and path planning in automated driving. Assuming that the precise vehicle position on the digital map is estimated, a region-of-interest (ROI) location for the traffic lights can be calculated using the registered traffic light position and current vehicle pose. Extracting the ROI makes it possible to reduce the search region of traffic lights. It is then possible to reduce false detections such as false-positive and false-negative detections, and computational costs [8, 9]. In addition to improving recognition performance, associating traffic lights registered on a map with traffic lights in an image is an important aspect of the map-based recognition. In the decision making using the HD maps, it is necessary to grasp the state of the relevant traffic lights, in order to make an approach decision at the intersection.

3.2 Related Works

Currently, many methods are proposed for traffic light recognition. Based on the state-of-the-art researches, the recognition procedure can be described as follows:

1. Determine the search region: A region-of-interest (ROI) is extracted from the captured image by using the predefined map.
2. Extract candidate objects: Circular lighting areas or rectangular objects are extracted from the search region as candidate traffic lights.
3. Classify the state of the candidates: Simple color filtering or machine learning algorithms identify lighting colors and arrow light directions.

Although different approaches have been developed, most of the methods involve extracting candidates according to their specific color spaces [10, 11] and circular shape [12, 13], and identifying them as traffic lights or arrow lights. Regardless of the country, the traffic lights mainly consist of circle and arrow shaped lights. In the case of the ROI-based recognition, detection of circular objects is one of the effective approaches to recognize lighting areas because the search region is limited in it. In literatures proposed up to now, some methods adopted a blob detector which extracted candidate objects by binarizing the image and segmenting pixels [10, 14, 15]. It can detect circular objects even if their size is a few pixels. Then, the recognition of the whole shape of the traffic lights are implemented using specific shape matching and machine learning. Moreover, the effect of introducing object tracking to stabilize the recognition result has been reported [10, 16]. In recent years, there have been reports of cases in which performance is improved upon using deep neural network (DNN) [10, 14, 17, 18, 19]. In order to detect arrow lights, machine learning-based detector is a key solution. It has been reported that these methods can recognize traffic lights at distances exceeding 100 m, with a recognition rate of approximately 90%.

On the other hand, assuming that algorithm will be introduced into the automated vehicles, real-time processing is important for decision making. In addition, in order to reduce the delay in recognition, it is necessary to recognize traffic lights in an appropriate processing time in accordance with the velocity of the vehicle. For example, when traveling at a velocity of 50 km/h, a vehicle moves about 14 m per second. Then, it is important to estimate the required time in consideration of the responsive deceleration for practical development.

3.3 Traffic Light Detection Algorithm

In this section, as an example of traffic light recognition algorithm, our previously proposed method [12] is briefly introduced.

In the traffic light recognition, the task is to recognize the state of the traffic lights in the image that corresponds to the HD map. As shown in Fig.2(a), (b), it is necessary to properly recognize the lighting status of traffic lights both in the day and the night. On the HD map, the position information of the traffic lights is recorded individually, and then the traffic light positions in the camera image can be calculated from the position of the traffic light and the vehicle. Fig.2(c) indicates the typical ROI image which is extracted by the coordinate transform using the HD map for a driving image. As in the enlarged image in Fig.2(c), the extracted ROI image may include traffic lights other than the ones, and background lighting objects. In implementing decision making algorithm at intersections for autonomous vehicle, the purpose is to recognize the traffic light associated with the ROI. Therefore, if a different traffic light is recognized in the specific ROI, it will be a false-positive detection.

![Fig 2 Typical traffic light image at different brightness and region of interest (ROI) image](image-url)
Fig. 3 illustrates a flowchart of our method. It mainly consists of the following five procedures:

1. Search target traffic lights and compute ROI in image using HD maps.
2. Generate a highlighted image as a feature image which emphasizes the lights of traffic lights in the ROI image.
3. Extract candidates for traffic light in the generated highlighted image using three types of different methods.
4. Compute the probability of existence area containing traffic lights using a time-series processing.
5. Recognize the arrow light, if the target traffic light has attribute information of arrow light.

As shown in Fig. 4(a), there is a situation that the occluded situation where it is difficult to see the whole shape of the traffic light. However, it is necessary to recognize the traffic light state only from the lighting area. The situation where such an overall shape cannot be visually recognized is the same even at night as shown in Fig. 4(b).

Arrow detection requires recognition of directions of arrows. An arrow signal is recognized when the traffic light has the attribute of arrow lights in the HD map. In Japanese traffic environment, arrow lights are generally lit at red and yellow traffic lights. After detecting a yellow or red signal, an arrow detection ROI is determined from detected traffic light position. In the arrow recognition process, the right-arrow detector is trained in advance using AdaBoost, and then it is applied to the extracted ROI. In order to detect left/straight arrows, the ROI image is rotated, and the same detector is used to search objects.

By using the method described above, the traffic light recognition is realized by detecting the candidate objects, classifying the lighting color, and computing the confidence using the existence probability. This work further improves the recognition performance, especially for distant arrow lights, by utilizing the prior information given in the digital map. In the traffic light recognition, when there are multiple candidate objects, it is possible to weight candidates according to the distance of the traffic lights in the probability updating procedure. It is expected to reduce false-positive detections in background.

On the other hand, in arrow recognition, recognition can be improved by providing the pattern of the target traffic light as prior information. For example, Fig. 5 illustrates the typical arrow recognition scene. In the recognition of a distant arrow light, if it is difficult to visually recognize the direction of the lighting arrow, it may cause false-positives.
or false-negatives. In Fig.5, it can be seen that some arrow lights are lit in the ROI image, but it is difficult to distinguish the directions. In this case, because the lighting parts of the arrow light is crushed, a candidate point may be detected at the arrow traffic light as well as the candidate traffic light. Normally, this detected candidate can be a false-positive detection of a green signal. However, if information on the relative positional relationship of the arrow lights at the traffic light is provided as a prior information, it is possible to distinguish the direction of the arrow lights.

4. Issues in Traffic Light Recognition

4.1 Evaluation of Traffic Light Detection Algorithm

As described above, as part of the SIP project, the authors' group is conducting test runs in the Tokyo waterfront area of Japan and is collecting a large amount of data using onboard sensors mounted on the test vehicles, including image data for traffic light recognition. In this section, we evaluate the method described in Section 3 using image data for traffic light recognition obtained in the test run.

![Fig. 6 Accuracy of traffic light recognition](image)

In this evaluation, in order to evaluate the situation close to the actual autonomous driving in the urban area, the image list at the approaching of a series of intersections is defined as a scenario, and a large number of scenario data are evaluated collectively. The evaluation data consisted of 233 intersection approach scenarios based on data from the daytime, backlight, nighttime and rainy weather conditions. The total number of frames of image data is 42,603, and the total number of traffic lights and ROI to be evaluated is 97,985. These included 81,273 traffic lights with only illuminated traffic signals, and 8,555 traffic lights that had red and arrow lights. Furthermore, there were 8,157 traffic lights where it was not possible to confirm whether the ROI signals were illuminated due to occlusion by vehicles or buildings. Using these data, the distance from the vehicles to traffic lights was divided into 10 m sections, and the mean F-value was calculated, and the recognition rate was evaluated. Fig.6 shows the recognition rates for blue, red, and arrow lights for each distance for all the evaluation data. Furthermore, as shown in Fig.6(b) and (d), in addition to evaluating the recognition performance in terms of individual traffic lights, it was also evaluated in terms of each intersection. This is because it is believed that when multiple traffic lights are installed in the same intersection, the state of traffic light can be determined through a majority vote strategy from the state of the multiple traffic lights, even if some cannot be seen due to occlusion by other objects. In addition, although the cameras currently in use have Full HD resolution, to evaluate how much the recognition rate increases if a camera of higher resolution is available in the future, the recognition rate when a camera with a telephoto lens in addition to normal lens was evaluated. The lens used here has field of view of 27deg and 53 deg. The evaluation results are shown in Fig.6(c) and (d).

From Fig.6 (a), when the recognition rate was evaluated in terms of individual traffic lights, it was confirmed that it especially drops for arrow lights located a long distance away, whereas those at a short distance could appropriately be recognized. This is because with currently used cameras, at above 90 m, for example, the number of pixels for an arrow light is 10 pixel or less, making it difficult to accurately recognize the direction of the arrow. In contrast, it was confirmed that long-distant recognition rate can be improved by integrating the results of multiple traffic lights at an intersection and/or the combined use of telephoto lens as shown in Fig.6(b), (c) and (d).

As listed in Table 1, it was also confirmed that approximately 99.0% of the traffic light within 120 m were recognizable.

<table>
<thead>
<tr>
<th>Recognition accuracy</th>
<th>For each Traffic light</th>
<th>For each Intersection</th>
<th>For each Intersection with telephoto lens</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue light</td>
<td>0.984</td>
<td>0.996</td>
<td>0.997</td>
</tr>
<tr>
<td>Red light</td>
<td>0.977</td>
<td>0.987</td>
<td>0.992</td>
</tr>
<tr>
<td>Arrow signal</td>
<td>0.908</td>
<td>0.960</td>
<td>0.982</td>
</tr>
<tr>
<td>Mean value</td>
<td>0.956</td>
<td>0.981</td>
<td>0.990</td>
</tr>
</tbody>
</table>

4.2 Issues in Traffic Light Recognition

In the evaluation in the previous section, it was shown that traffic lights were approximately 99% recognizable within 120 m, although they were evaluated in a limited area of the Tokyo waterfront area of Japan. In this section, we consider situations in which the recognition of traffic lights becomes difficult with current technology based on the results of the evaluations.

Fig.7 shows the four characteristic situations in traffic light recognition evaluation results in which recognition was difficult, such as occlusion, background assimilation,
nighttime, and backlight. Fig.7(a) shows a scene where traffic lights were occluded by roadside trees extending into the road. Additionally, there were scenes in which traffic lights were not visible in the traffic environment due to large vehicles or roadside structures. In situations where traffic lights cannot physically be seen, it becomes difficult to maintain reliable decision making at intersection from only from image sensors. Fig.7(b) shows an example in which the boundary of a traffic light became unclear and undetectable due to the building in the background surrounding the traffic light. Furthermore, Fig.7(c) shows an example in which recognition became difficult at night. In the nighttime image, the outline around the illuminated object tended to be blurred in comparison to daytime due to the effects of glare and blur around the illuminated object. Accordingly, as shown in Fig.7(c), problems are confirmed such as undetected arrow lights or mis-detection of blue traffic lights because of the brightness saturation in the background region caused by sunlight. In this case, because it became impossible to properly recognize traffic lights in the vicinity of the saturated area, the issues of missing or incorrect detection were confirmed. Furthermore, as can be confirmed in Fig.7(d), the saturated area may have a hue close to yellow on the captured image, and cases have occurred in which a red light or a blue light is incorrectly recognized as a yellow signal. However, although it is difficult to recognize these problems at long distances such as in case of background assimilation and at nighttime scene, there are many cases where recognition can be performed at short distances without problems. Although it became temporarily difficult to recognize a few frames in the middle of approaching intersections, there were no scenarios in the evaluation results in which this fatally impacted decision-making to enter the intersections.

However, although not present in the situations evaluated here, it is expected that if a situation occur in which a traffic light cannot be recognized due to the effects of backlight, for example, when stopped at the start of an intersection, deciding whether to enter an intersection will be difficult. Based on the results of our basic experiments, it is estimated that when sunlight is present near the traffic lights in images, the area saturated with brightness as a result of this effect is evaluated to be approximately 5 to 10 degrees. Therefore, if multiple traffic lights are installed at the intersection for example, it is considered that it will be possible to decide whether to enter an intersection if other traffic lights are visually recognizable. On the other hand, if only one traffic light is installed at the intersection, or if all the traffic lights cannot be visually recognized due to the influence of sunlight even if multiple traffic lights are installed, it may be difficult to judge the approach to the intersection. Especially in the case of lamp-type traffic lights that have been installed in Japan for a long time, there are situations where all traffic lights seems to be lit as shown in Fig.8. This situation may be occurred front light situation where sunlight is present in front direction of the traffic light. As a result of basic experiments, it was found that the range of this effect is wide. Since all traffic lights in the Tokyo waterfront area were the LED-type, such effects were not apparent. However, because many places in Japan still have many lamp-type traffic lights, there is a concern that multiple traffic lights may not be visible at the same time due to the front light issue.

For this reason, it turned out that infrastructure measures are also important in addition to improving recognition technology, such as installing multiple traffic lights at an intersection, and devising the arrangement of traffic lights so that the effects of backlight, etc. do not occur at the same time. In addition, it is considered that replacing older lamp-type traffic lights with LED-type traffic lights will also be important for ensuring traffic light recognition.

![Image](blue_light_illuminated.png)

Fig.8 View of lamp-type traffic light with front light (Blue light is illuminated)
5. Summary

In this paper, we discussed the issues facing traffic light recognition technology, using traffic light recognition algorithm developed by the authors as an example. From the evaluation results using the onboard sensor data from autonomous driving currently being conducted by the authors in the Tokyo waterfront area, it was proven that traffic lights could be recognized with an accuracy of approximately 99%.

However, the evaluation results confirmed that traffic light recognition is difficult under conditions such as "occlusion", "background assimilation," "nighttime," and "backlight." It was also confirmed that these effects are mostly temporary, and do not significantly affect decision-making to enter an intersection, because information from multiple traffic lights installed in intersections can be used. However, in the situation where all the traffic lights cannot be visually recognized because of the effects of backlight or front light while stopped at the start of an intersection, recognition using current onboard cameras is expected to become technically difficult. Therefore, it is inferred that there is a need to explore how to properly install traffic lights with the assumption of recognition by on-board cameras.

On the other hand, since autonomous driving systems require reliable recognition and decision making, it is also desirable to improve reliability by designing redundant systems. Although, the one idea to create redundant system is to install multiple onboard cameras, it should be carefully considered that issues caused by light problem may not be solved because theoretically possible that cause of the problem will be same if we use same principal of sensors. Therefore, for example, it is also desirable to install a traffic light equipped with communication equipment.

In preparation for the coming era of autonomous driving, it is necessary that we should simply improve recognition technology using onboard sensors, but also accelerate discussions on how infrastructure should be developed from the perspective of ease of recognition. It is believed that this will make it possible to improve the reliability of the autonomous driving system, and it will also contribute to improving the safety of human drivers simultaneously.

Acknowledgments

This work was supported by Council for Science, Technology and Innovation (CSTI), Cross-ministerial Strategic Innovation Promotion Program (SIP), “Research and development for recognition technologies and other technologies necessary for automated driving technologies (levels 3 and 4)” (funded by NEDO).

References

2018.


Naoki Suganuma received the B.E., M.E., and Ph.D. degrees from Kanazawa University in 1998, 2000, 2002, respectively. Since 2002, he is with the Kanazawa University, and currently he is a professor and vise director of advanced mobility research institute of Kanazawa University. His research interests are autonomous vehicle. He is a member of JSAE, SICE, RSJ, IATS and, JSME.

Keisuke Yoneda is currently working as an Associate Professor at Automated Vehicle Research Unit of Institute for Frontier Science Initiative, Kanazawa University. He received his B.E. degrees in information engineering from Toyohashi University of Technology in 2007, and M.S., Ph.D. degrees in information science from Hokkaido University in 2009 and 2012, respectively. He was working as a Postdoctoral Fellow at Toyota Technological Institute (2012-2015), and an Assistant Professor at Kanazawa University (2015-2020). His research interests are autonomous vehicles, artificial intelligence, and computer vision. He is a member of IEEE and the Society of Automotive Engineers of Japan, Inc.